GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 17, 2003, 08:44:13 ; Search time 42 Seconds
(without alignments)
2468.309 Million cell updates/sec

Title: US-10-010-227-3
Perfect score: 4055
Sequence: 1 MPGAESTPQTLYDKVLQAHV.....KAVPVPTTNRGEEKKEPLEW 778

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_101002:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666.5	41.1	644	22	AAU34234	Staphylococcus aur
2	1507.5	37.2	474	22	AAU33631	Pseudomonas aerugi
3	1479.5	36.5	466	22	AAU34428	E. coli cellular p
4	1471	36.3	469	22	AAU35571	Haemophilus influe
5	1450.5	35.8	466	22	AAU38424	Salmonella typhi c
6	1347.5	33.2	481	22	AAG91198	C glutamicum prote
7	1330	32.8	553	22	AAB79766	Corynebacterium gl
8	1316	32.5	534	22	AAB79767	Corynebacterium gl
9	1300	32.1	461	23	ABP39047	Staphylococcus epi
10	1282	31.6	456	22	AAU36563	Staphylococcus aur
11	1282	31.6	456	22	AAU37335	Staphylococcus aur
12	1279.5	31.6	462	23	ABB48172	Listeria monocytog
13	1138.5	28.1	460	15	AAR54216	L.lactis branched
14	1132.5	27.9	460	23	ABB54552	Lactococcus lactis
15	749	18.5	264	19	AAW77717	3-Isopropylmalate
16	688.5	17.0	245	22	AAG81974	S. epidermidis ope
17	637.5	15.7	423	22	AAB96358	Putative 3-isoprop
18	558.5	13.8	461	21	AAG29924	Arabidopsis thalia
19	558.5	13.8	469	21	AAG40226	Arabidopsis thalia
20	558.5	13.8	509	21	AAG29923	Arabidopsis thalia
21	558.5	13.8	509	23	ABB92924	Herbicidally activ
22	558.5	13.8	517	21	AAG40225	Arabidopsis thalia
23	550.5	13.6	443	21	AAG29925	Arabidopsis thalia
24	550.5	13.6	451	21	AAG40227	Arabidopsis thalia
25	542	13.4	212	22	AAU36307	Pseudomonas aerugi
26	528.5	13.0	428	21	AAY28433	Corn leuC subunit
27	523.5	12.9	418	21	AAB15394	Thermus thermophil
28	523.5	12.9	443	21	AAY28434	Corn leuC subunit
29	506	12.5	201	22	AAU38423	Salmonella typhi c
30	495	12.2	200	22	AAU35572	Haemophilus influe
31	493	12.2	173	19	AAW77716	3-Isopropylmalate
32	491	12.1	201	22	AAU34427	E. coli cellular p
33	482.5	11.9	191	15	AAR54217	L.lactis branched
34	482.5	11.9	191	23	ABB54553	Lactococcus lactis
35	482.5	11.9	780	22	AAB84333	Amino acid sequenc
36	481.5	11.9	780	20	AAW98947	Mitochondrial acon
37	472.5	11.7	787	22	ABB58512	Drosophila melanog
38	464.5	11.5	146	19	AAW79390	Staphylococcus aur
39	461.5	11.4	200	22	AAU36055	Klebsiella pneumon
40	442	10.9	683	22	ABB58511	Drosophila melanog
41	424	10.5	769	22	ABB60903	Drosophila melanog
42	422	10.4	190	22	AAU36564	Staphylococcus aur
43	403	9.9	732	22	ABG05728	Novel human diagno
44	396	9.8	189	23	ABP39041	Staphylococcus epi
45	393	9.7	899	22	ABB62277	Drosophila melanog
46	389.5	9.6	197	22	AAG91199	C glutamicum prote
47	389.5	9.6	197	22	AAB79772	Corynebacterium gl
48	389	9.6	236	21	AAB41138	Human ORFX ORF902
49	380.5	9.4	902	22	ABB61080	Drosophila melanog
50	375.5	9.3	193	23	ABB48173	Listeria monocytog
51	375.5	9.3	304	21	AAG27662	Arabidopsis thalia
52	375.5	9.3	901	22	AAB83182	Corynebacterium th
53	370.5	9.1	291	21	AAG27663	Arabidopsis thalia
54	367.5	9.1	943	22	AAG91445	C glutamicum prote
55	366.5	9.0	805	22	AAU62827	Propionibacterium
56	365.5	9.0	868	21	AAY93301	Amino acid sequenc
57	364.5	9.0	901	22	AAG81624	S. epidermidis ope
58	364.5	9.0	914	23	ABP40363	Staphylococcus epi
59	363	9.0	919	16	AAR84337	Arabidopsis thalia
60	363	9.0	927	16	AAR85598	Arabidopsis thalia
61	359.5	8.9	869	21	AAY93285	Amino acid sequenc
62	357	8.8	891	22	AAU29353	Novel mar regulate
63	356.5	8.8	380	22	AAB96131	Putative 3-isoprop
64	354.5	8.7	889	23	ABB57134	Mouse ischaemic co
65	353.5	8.7	764	16	AAR84345	Melon aconitase en
66	353	8.7	164	18	AAW28261	Amino acid sequenc
67	352.5	8.7	891	16	AAR84338	Maize aconitase pa
68	344.5	8.5	802	22	ABG22298	Novel human diagno
69	340	8.4	160	22	AAG81975	S. epidermidis ope
70	336	8.3	110	23	ABP06700	Human ORFX protein
71	330	8.1	144	22	AAU53144	Propionibacterium
72	329	8.1	963	23	AAE19852	Rat wild-type IRP-
73	328.5	8.1	963	22	AAU32389	Novel human secret
74	326.5	8.1	952	23	AAE19851	Human wild-type IR
75	326	8.0	117	22	AAU53143	Propionibacterium
76	325	8.0	848	23	ABB53982	Lactococcus lactis
77	324	8.0	900	23	ABB48185	Listeria monocytog
78	318.5	7.9	106	23	ABP32040	Human isomerase-li
79	312.5	7.7	139	23	ABP32235	Human isomerase-li
80	309.5	7.6	639	21	AAY74439	Neisseria meningit
81	307	7.6	639	21	AAY74440	Neisseria meningit
82	303	7.5	852	22	AAU35942	Helicobacter pylor
83	303	7.5	853	22	AAU35770	Helicobacter pylor

84	301	7.4	639	21	AAY74438	Neisseria gonorrhe
85	292	7.2	86	23	ABP08386	Human ORFX protein
86	284.5	7.0	214	21	AAG27664	Arabidopsis thalia
87	271	6.7	814	22	ABG19229	Novel human diagno
88	264.5	6.5	120	22	AAG81461	S. epidermidis ope
89	250	6.2	127	21	AAY28435	Rice leuC subunit
90	239	5.9	119	19	AAW38591	S. pneumoniae isop
91	238	5.9	70	23	ABP04569	Human ORFX protein
92	212.5	5.2	268	18	AAW20584	H. pylori cytoplas
93	209	5.2	78	19	AAW79391	Staphylococcus aur
94	209	5.2	263	21	AAY28441	LeuD subunit of 3-
95	208.5	5.1	195	21	AAY28436	LeuC subunit of 3-
96	207.5	5.1	533	23	ABB89994	Human polypeptide
97	207	5.1	205	22	ABG05731	Novel human diagno
98	204.5	5.0	553	22	AAB79520	Corynebacterium gl
99	204.5	5.0	557	22	AAB79519	Corynebacterium gl
100	198.5	4.9	203	18	AAW20430	H. pylori cytoplas

ALIGNMENTS

```
RESULT 1
AAU34234
ID   AAU34234 standard; Protein; 644 AA.
XX
AC   AAU34234;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #510.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS52093.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 5730; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   644 AA;

  Query Match             41.1%;  Score 1666.5;  DB 22;  Length 644;
  Best Local Similarity   48.7%;  Pred. No. 2.3e-139;
  Matches  352;  Conservative   92;  Mismatches  198;  Indels   81;  Gaps   10;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           |||:|||   ||:  ||    ||||| ||:||||||||||||||   ||:|||| | || |
Db       3 QTLFDKVWNRHVLYGKLGEPQLLYIDLHLIHEVTSPQAFEGLRLQNRKLRRPDLTFATLD 62

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
           |||||       || : ||:: :  |  ||::|  :|||  | :    |||||::||| |
Db      63 HNVPTI------DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDEQGIVHMVGPETG 115

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           | || |:|||||||:|||||||:|||||||||||| ||| |   : ||::| ::| ||
Db     116 LTQPGKTIVCGDSHTATHGAFGAIAFGIGTSEVEHVFATQTLWQTKPKNLKIDINGTLPT 175

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           || :||::|| |   |   |||  :|| |  |::|||: ||:||||:|||||: |:: ||
Db     176 GVYAKDIILHLIKTYGVDFGTGYALEFTGETIKNLSMDGRMTICNMAIEGGAKYGIIQPD 235

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           :|||||:|||| |  :      |:  |: | ||  | :|  : :|   : | :||||:|
Db     236 DITFEYVKGRPFADNF-----AKSVDKWRELYSDDDAIFDRVIELDVSTLEPQVTWGTNP 290

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
           |  |  :   |       |    : :|   ||||: |   ||| :  ||:|||||:|: |
Db     291 EMGVNFSEPFP-------EISDINDQRAYDYMGLEPGQKAEDIDLGYVFLGSCTNARLSD 343

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
           |  |: :||| |  ||: :|:|||||  || :||: ||| ||: ||||||| ||||||||
Db     344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKREAEKLGLDTIFKNAGFEWREPGCSMCLGM 402

Qy     429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKLTDYKASPH 488
           ||| :     |||||||||||||| | ||||:|| ||||| | ||  ||||:
Db     403 NPDQVPEGVHCASTSNRNFEGRQGKGARTHLVSPAMAAAAAAIHGKFVDVRKV-------- 454

Qy     489 IAAYQKSTVTKPHVDERINQDAHEKDIIADIPEDNNGPHTNTSASVGTSAGLPKFTILKG 548
                                                 |   | :   |  ||
Db     455 ----------------------------------------------VVXMAAIKPITTYKG 469

Qy     549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDFVLNKEPYRKAS 608
              ||   |:||| ||||  || | ::| |   | | |:  ||::  ||  ||  |: ||
Db     470 KIVPLFNDNIDTDQIIPKVHLKRISKSGFGPFAFDEWRYLPDGSDNPDFNPNKPQYKGAS 529

Qy     609 ILVCTGANFGCGSSREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGMLPIPIKDQAQIE 668
           ||: || |||||||||| ||| |:|   :|| ||:|||: |  || |||| ::  |: |
Db     530 ILI-TGDNFGCGSSREHAAWALKDYGFHIIIAGSFSDIFYMNCTKNAMLPIVLEKNAR-E 587

Qy     669 AIAAEARAGKEIEVDLPNQLIKNATGETICTFEVEEFRKHCLVNGLDDIGLTMQMEDKIA 728
           :|       |||||||| :  :: :    ||::|  |: |||||||| :|:| |  |
Db     588 HLAKYV----EIEVDLPNQTV--SSPDKSFHFEIDETWKNKLVNGLDDIAITLQYESLIE 641

Qy     729 EFE 731
           ::|
Db     642 KYE 644


RESULT 2
AAU33631
ID   AAU33631 standard; Protein; 474 AA.
```

```
XX
AC   AAU33631;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa cellular proliferation protein #75.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS51490.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 5127; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   474 AA;

  Query Match           37.2%;  Score 1507.5;  DB 22;  Length 474;
  Best Local Similarity   62.5%;  Pred. No. 2.1e-125;
  Matches  295;  Conservative   61;  Mismatches  113;  Indels    3;  Gaps    2;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           :||||||:   |:| :: ||: |:|||||::|||||||||||||| ||||  | |  :|| |
Db       4 KTLYDKLWDMHLVKQRDDGSALIYIDRHILHEVTSPQAFEGLRLAGRKPWRIDANIATPD 63

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
           ||||||  :    ||: | :: || |  ||:||  :||:| | ::| |||||||:|||||
Db      64 HNVPTTRTERKGGIAA-IADEVSRLQVQTLDENCDDFGITEFKMNDVRQGIVHVVGPEQG 122

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           ||||  |||||||||||||||||| ||||||||||||||: |: ||| ::|:| |
Db     123 ATLPGMTVVCGDSHTSTHGAFGALAHGIGTSEVEHVLATQCLVAKKMKNMLVKVEGRLPA 182

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           ||::||:||  || |||||| |  ||| || || ||:| ||:||||||| ||| |:|| |
Db     183 GVTAKDIVLAVIGRIGTAGGNGHAIEFAGSAIRDLSIEGRMTICNMSIEAGARVGLVAVD 242

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           : | :|:|||| ||   : :| :|   |: | ||  |::|  | :||  | | ::|||||
Db     243 QKTIDYVKGRPFAP--SAEQWDQAVACWQGLVSDADARFDTVVELDAAQIKPQVSWGTSP 300

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
           | |: :   ||||   :   |:    | |:||||:    : || :|:||||||||||||
Db     301 EMVLAVDQNVPDPARESDPIKRGSIERALKYMGLRPNQAITDIQLDRVFIGSCTNSRIED 360

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
           ||||| | :||| |  :| |:|||||||| |||:||||:|| |||||||| |||||| |
Db     361 LRAAAEVARGRKVAATIKQALVVPGSGLVKEQAEKEGLDRIFIEAGFEWREPGCSMCLAM 420

Qy     429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
           ||| |   | ||||||||||||||||||:|| ||||| : |:  |||:|
Db     421 NPDRLESGEHCASTSNRNFEGRQGAGGRTHLVSPAMAAAAAVNGRFIDVREL 472


RESULT 3
AAU34428
ID   AAU34428 standard; Protein; 466 AA.
XX
AC   AAU34428;
XX
DT   14-FEB-2002  (first entry)
XX
DE   E. coli cellular proliferation protein #9.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Escherichia coli.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS52287.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 10021; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
```

```
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   466 AA;

  Query Match           36.5%;  Score 1479.5;  DB 22;  Length 466;
  Best Local Similarity   62.1%;  Pred. No. 6.6e-123;
  Matches  293;  Conservative   59;  Mismatches  111;  Indels    9;  Gaps    4;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
Db       3 KTLYEKLFDAHVVYEAENETPLLYIDRHLVHEVTSPQAFDGLRAHGRPVRQPGKTFATMD 62

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
Db      63 HNVSTQT----KDINAC--GEMARIQMQELIKNCKEFGVELYDLNHPYQGIVHVMGPEQG 116

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
Db     117 VTLPGMTIVCGDSHTATHGAFGALAFGIGTSEVEHVLATQTLKQGRAKTMKIEVQGKAAP 176

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
Db     177 GITAKDIVLAIIGKTGSAGGTGHVVEFCGEAIRDLSMEGRMTLCNMAIEMGAKAGLVAPD 236

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
Db     237 ETTFNYVKGRLHAPK--GKDFDDAVAYWKTLQTDEGATFDTVVTLQAEEISPQVTWGTNP 294

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
Db     295 GQVISVNDNIPDPASFADPVERASAEKALAYMGLKPGIPLTEVAIDKVFIGSCTNSRIED 354

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
Db     355 LRAAAEIAKGRKVAPGVQ-ALVVPGSGPVKAQAEAEGLDKIFIEAGFEWRLPGCSMCLAM 413

Qy     429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
Db     414 NNDRLNPGERCASTSNRNFEGRQGRGGRTHLVSPAMAAAAAVTGHFADIRNI 465


RESULT 4
AAU35571
ID   AAU35571 standard; Protein; 469 AA.
XX
AC   AAU35571;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Haemophilus influenzae cellular proliferation protein #212.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Haemophilus influenzae.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS53430.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 11164; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   469 AA;

  Query Match           36.3%;  Score 1471;  DB 22;  Length 469;
  Best Local Similarity   59.4%;  Pred. No. 3.8e-122;
  Matches  281;  Conservative   72;  Mismatches  112;  Indels    8;  Gaps    2;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
Db       3 KTLYEKLFDSHIVYEAEGETPILYINRHLIHEVTSPQAFDGLRVANRQVRQVNKTFGTMD 62

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
Db      63 HSISTQVRDVNK------LEGQAKIQVLELDKNTKATGIKLFDITTKEQGIVHVMGPEQG 116

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
Db     117 LTLPGMTIVCGDSHTATHGAFGALAFGIGTSEVEHVLATQTLKQARAKSMKIEVRGKVAS 176

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
Db     177 GITAKDIILAIIGKTTMAGGTGHVVEFCGEAIQDLSMEGRMTVCNMAIEMGAKAGLIAPD 236

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
Db     237 ETTFAYLKDRPHAPK--GKDWEDAVAYWKTLKSDDDAQFDTVVTLEAKDIAPQVTWGTNP 294

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
Db     295 GQVISVNETIPNPQEMADPVQRASAEKALHYIGLEAGTNLKDIKVDQVFIGSCTNSRIED 354

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
```

```
          |||||||:||||||  |||  :|||||||||  |||:|||||||  || |||  ||||||||
Db    355 LRAAAAVMKGRKKADNVKRILVVPGSGLVKEQAEKEGLDKIFIAAGAEWRNPGCSMCLGM 414

Qy    429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKLT 481
          | | |   ||||||||||||||||  |||||:|| ||||||: ||  |:| :|
Db    415 NDDRLGEWERCASTSNRNFEGRQGRNGRTHLVSPAMAAAAGVFGKFVDIRDVT 467


RESULT 5
AAU38424
ID   AAU38424 standard; Protein; 466 AA.
XX
AC   AAU38424;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Salmonella typhi cellular proliferation protein #315.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Salmonella typhi.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS56283.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 14017; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   466 AA;

  Query Match             35.8%;  Score 1450.5;  DB 22;  Length 466;
  Best Local Similarity   60.6%;  Pred. No. 2.5e-120;
  Matches  286;  Conservative   63;  Mismatches  114;  Indels    9;  Gaps    4;

Qy      9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
          :|||:|:  |||| |  : | |||||||||||||||||:|||   | ||:|  | || |
Db      3 KTLYEKLFDAHVVFEAPNETPLLYIDRHLVHEVTSPQAFDGLRAHHRPVRQPGKTFATMD 62

Qy     69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
          ||| | :    ||| :    : :| |   | :|  ||||  : |:    |||||:||||
Db     63 HNVSTQT----KDINA--SGEMARIQMQELIKNCNEFGVELYDLNHPYXGIVHVMGPEQX 116

Qy    129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
          ||||  :|||||||: ||||||| |||||||||||||| |   |:| |:|:| |  ||
Db    117 VTLPGMXIVCGDSHTAXHGAFGALXFGIGTSEVEHVLATQTLKQGRAKTMKIEVTGNAAP 176

Qy    189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
          |:::||:||  ||  |:||||| |:|||| ||:|||| ||::|||:|| ||:||:||||
Db    177 GITAKDIVLAIIGKTGSAGGTGHVVEFCGDAIRALSMEGRMTLCNMAIEMGAKAGLVAPD 236

Qy    249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
          | || |:||| |||    :: :| :||| |::| || :|  | : |::| | :||||:|
Db    237 ETTFNYVKGRLHAPK--GRDFDEAVEYWKTLKTDDGATFDTVVTLRAEEIAPQVTWGTNP 294

Qy    309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
            |: :| ::||| :|:   ::|   : | ||||: | |: |: :|||||||||||||||
Db    295 GQVISVTDIIPDPASFSDPVERASAEKALAYMGLQPGVPLTDVAIDKVFIGSCTNSRIED 354

Qy    369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
          ||||| | |||| || |: |:|||||| || ||| ||||||| |||||||  |||||| |
Db    355 LRAAAEVAKGRKVAPGVQ-ALVVPGSGPVKAQAEAEGLDKIFIEAGFEWRLPGCSMCLAM 413

Qy    429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
          | | | | ||||||||||||||||  |||||:|| ||||| : |  ||:| :
Db    414 NNDRLNPGERCASTSNRNFEGRQGRGGRTHLVSPAMAAAAAVTGHFADIRSI 465


RESULT 6
AAG91198
ID   AAG91198 standard; Protein; 481 AA.
XX
AC   AAG91198;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 4952.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   N-PSDB; AAH66417.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
```

XX
PS Claim 17; SEQ ID NO: 4952; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 481 AA;

Query Match 33.2%; Score 1347.5; DB 22; Length 481;
Best Local Similarity 57.6%; Pred. No. 4.1e-111;
Matches 273; Conservative 61; Mismatches 127; Indels 13; Gaps 6;

```
Qy      10 TLYDKVLQAHVVDEKLDGTV-LLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           || :|| : ||| :  :|    ||||| |:||||||||:|||  |||:| |:  ||| |
Db      15 TLAEKVWRDHVVSKGENGEPDLLYIDLQLLHEVTSPQAFDGLRMTGRKLRHPELHLATED 74

Qy      69 HNVPTTSRKALKDIASFIKEDD--SRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPE 126
           |||||   |      | :: :|  || |  || :| :||||    : | |||||| :||:
Db      75 HNVPTEGIKT----GSLLEINDKISRLQVSTLRDNCEEFGVRLHPMGDVRQGIVHTVGPQ 130

Qy     127 QGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGEL 186
           | | || |:|||||||||||||::|||||||||||:||| |   |   | | |:| |||
Db     131 LGATQPGMTIVCGDSHTSTHGAFGSMAFGIGTSEVEHVMATQTLPLKPFKTMAIEVTGEL 190

Qy     187 APGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVA 246
           |||||||::|  |  ||| || | |:|: |  || :||:|||::|||||| ||||||:|
Db     191 QPGVSSKDLILAIIAKIGTGGGQGYVLEYRGEAIRKMSMDARMTMCNMSIEAGARAGMIA 250

Qy     247 PDEITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGT 306
           ||: ||:|::|| :|||    :| :|  ||| | :| || :|  | ||   : | :||||
Db     251 PDQTTFDYVEGREMAPK--GADWDEAVAYWKTLPTDEGATFDKVVEIDGSALTPFITWGT 308

Qy     307 SPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRI 366
           :|   :|:   || || |  :  ||   : |||| |  |||: || :| ||:|||||:||
Db     309 NPGQGLPLGESVPSPEDFTNDNDKAAAEKALQYMDLVPGTPLRDIKIDTVFLGSCTNARI 368

Qy     367 EDLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCL 426
           |||: || ::|| | |  ::  |||| |  :| :||  |||||| :|| ||| |||||||
Db     369 EDLQIAADILKGHKIADGMR-MMVVPSSTWIKQEAEALGLDKIFTDAGAEWRTAGCSMCL 427

Qy     427 GMNPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKL---ADV 477
           ||||| | | || |||||||||||||| ||||||:|| :|||  | | |   ||:
Db     428 GMNPDQLKPGERSASTSNRNFEGRQGPGGRTHLVSPAVAAATAIRGTLSSPADI 481
```

RESULT 7
AAB79766
ID AAB79766 standard; Protein; 553 AA.
XX
AC AAB79766;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:266.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF71885.
XX

PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 545-547; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 553 AA;

Query Match 32.8%; Score 1330; DB 22; Length 553;
Best Local Similarity 57.9%; Pred. No. 1.9e-109;
Matches 267; Conservative 60; Mismatches 124; Indels 10; Gaps 5;

```
Qy      10 TLYDKVLQAHVVDEKLDGTV-LLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           || :|| : ||| :  :|   ||||| |:||||||||:|||  |||:| |:  ||| |
Db      15 TLAEKVWRDHVVSKGENGEPDLLYIDLQLLHEVTSPQAFDGLRMTGRKLRHPELHLATED 74

Qy      69 HNVPTTSRKALKDIASFIKEDD--SRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPE 126
           |||||   |     | :: :|  || |  || :| :||||    : | |||||| :||:
Db      75 HNVPTEGIKT----GSLLEINDKISRLQVSTLRDNCEEFGVRLHPMGDVRQGIVHTVGPQ 130

Qy     127 QGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGEL 186
           | | || |:|||||||||||||::|||||||||||||:||| |   |  | | |:| |||
Db     131 LGATQPGMTIVCGDSHTSTHGAFGSMAFGIGTSEVEHVMATQTLPLKPFKTMAIEVTGEL 190

Qy     187 APGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVA 246
           |||||||::|  |  ||| || | |:|: |  || :||:|||::|||||| ||||||:|
Db     191 QPGVSSKDLILAIIAKIGTGGGQGYVLEYRGEAIRKMSMDARMTMCNMSIEAGARAGMIA 250

Qy     247 PDEITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGT 306
           ||: ||:|::|| :|||    :| :|  ||| | :| || :|  | ||   : | :||||
Db     251 PDQTTFDYVEGREMAPK--GADWDEAVAYWKTLPTDEGATFDKVVEIDGSALTPFITWGT 308

Qy     307 SPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRI 366
           :|   :|:   || || |  :  ||   : |||| |  |||: || :| ||:|||||:||
Db     309 NPGQGLPLGESVPSPEDFTNDNDKAAAEKALQYMDLVPGTPLRDIKIDTVFLGSCTNARI 368

Qy     367 EDLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCL 426
           |||: || ::|| | |  ::  |||| |  :| :||  |||||| :|| ||| |||||||
Db     369 EDLQIAADILKGHKIADGMR-MMVVPSSTWIKQEAEALGLDKIFTDAGAEWRTAGCSMCL 427

Qy     427 GMNPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAA 467
           ||||| | | || | |||||||||||  |||||:|| :|||
Db     428 GMNPDQLKPGERSAFTSNRNFEGRQGPGGRTHLVSPAVAAA 468
```

RESULT 8
AAB79767
ID AAB79767 standard; Protein; 534 AA.
XX
AC AAB79767;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:268.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF71886.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in

```
PT   microorganisms, including organic acids, nonproteinogenic amino acids,
PT   and purine and pyrimidine bases -
XX
PS   Claim 20; Page 549-551; 1737pp; English.
XX
CC   AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC   pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC   MP nucleic acids are useful for the production of fine chemicals
CC   in microorganisms, including organic acids, nonproteinogenic amino
CC   acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC   saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC   compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ   Sequence   534 AA;

  Query Match             32.5%;  Score 1316;  DB 22;  Length 534;
  Best Local Similarity   57.9%;  Pred. No. 3.1e-108;
  Matches  264;  Conservative   59;  Mismatches  123;  Indels   10;  Gaps    5;

Qy      15 VLQAHVVDEKLDGTV-LLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTDHNVPT 73
           | : ||| :  :|   ||||| |:||||||||||:|||  |||:| |:  ||| ||||||
Db       1 VWRDHVVSKGENGEPDLLYIDLQLLHEVTSPQAFDGLRMTGRKLRHPELHLATEDHNVPT 60

Qy      74 TSRKALKDIASFIKEDD--SRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQGFTL 131
             |      | :: :|  || |  || :| :||||    : | |||||| :||: | |
Db      61 EGIKT----GSLLEINDKISRLQVSTLRDNCEEFGVRLHPMGDVRQGIVHTVGPQLGATQ 116

Qy     132 PGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAPGVS 191
           || |:|||||||||||::||||||||||||:||| |  |  | | |:| ||| ||||
Db     117 PGMTIVCGDSHTSTHGAFGSMAFGIGTSEVEHVMATQTLPLKPFKTMAIEVTGELQPGVS 176

Qy     192 SKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPDEIT 251
           |||::|  |  ||| || | |:|: |  || :||:|||::|||||| ||||||:|||: |
Db     177 SKDLILAIIAKIGTGGGQGYVLEYRGEAIRKMSMDARMTMCNMSIEAGARAGMIAPDQTT 236

Qy     252 FEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSPEDV 311
           |:|::|| :|||    :| :|  ||| | :| || :|  | ||   : | :||||:|
Db     237 FDYVEGREMAPK--GADWDEAVAYWKTLPTDEGATFDKVVEIDGSALTPFITWGTNPGQG 294

Qy     312 VPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIEDLRA 371
           :|:   || || |  :  ||   : |||| |  |||: || :| ||:|||||:|||||:
Db     295 LPLGESVPSPEDFTNDNDKAAAEKALQYMDLVPGTPLRDIKIDTVFLGSCTNARIEDLQI 354

Qy     372 AAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGMNPD 431
           || ::|| | |  ::  |||| |  :| :||  |||||| :|| ||| ||||||||||
Db     355 AADILKGHKIADGMR-MMVVPSSTWIKQEAEALGLDKIFTDAGAEWRTAGCSMCLGMNPD 413

Qy     432 ILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAA 467
           | | || | |||||||||| ||||||:|| :|||
Db     414 QLKPGERSAFTSNRNFEGRQGPGGRTHLVSPAVAAA 449


RESULT 9
ABP39047
ID   ABP39047 standard; Protein; 461 AA.
XX
AC   ABP39047;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3892.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN91592.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 3892; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence   461 AA;

  Query Match             32.1%;  Score 1300;  DB 23;  Length 461;
  Best Local Similarity   54.7%;  Pred. No. 6.6e-107;
  Matches  258;  Conservative   69;  Mismatches  125;  Indels   20;  Gaps    5;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           |||:||| : ||:  |     ||||| ||:|||||||||||||   ||:|||| | || |
Db       8 QTLFDKVWKKHVLHGKEGEPQLLYIDLHLIHEVTSPQAFEGLRIQNRKLRRPDLTFATLD 67

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
           |||||       || : ||:: :  |  ||::| |:|||  | :    |||||::||| |
Db      68 HNVPTI------DIFN-IKDEIAHKQITTLQQNAKDFGVHIFDMGSDEQGIVHMVGPETG 120

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           | || |:|||||||:|||||||:||||||||||||| |||  |   : ||::| ::| |
Db     121 LTQPGKTIVCGDSHTATHGAFGAIAFGIGTSEVEHVFATQTLWQTKPKNLKININGSLPT 180

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           || :||::|: |   |   |||  :|| |  |::|||||||:||||:|| ||: |:: ||
Db     181 GVYAKDIILYLINQYGVDFGTGYALEFTGETIKNLSMEARMTICNMAIEAGAKYGLMQPD 240

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           | ||:|:|||| |  :||     :  :|| | ||  | :|  : :|  :: | :||||:|
Db     241 ETTFDYVKGRPYATDFDS-----SMAWWKKLYSDDDAYFDKVIELDVTNLEPQVTWGTNP 295

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
           |  |  :   |       | | |: :|   |||  |   ||| :  ||:|||||:|: |
Db     296 EMGVSFSNPFP-------EIKNANDQRAYDYMGLHPGQKAEDIKLGYVFLGSCTNARLSD 348

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
           |  |: ::||::  ||: :|:|||||  || :||  ||||:|::|||||| ||||||||
Db     349 LIEASHIIKGQQVHPNI-TAIVVPGSRTVKKEAEALGLDKLFKDAGFEWREPGCSMCLGM 407

Qy     429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
           ||| :     |||||||||||||| | ||||:|| ||||| | ||  ||||:
Db     408 NPDQVPEGVHCASTSNRNFEGRQGKGARTHLVSPAMAAAAAAINGKFIDVRKV 459


RESULT 10
AAU36563
ID   AAU36563 standard; Protein; 456 AA.
XX
AC   AAU36563;
XX
DT   14-FEB-2002  (first entry)
```

```
XX
DE   Staphylococcus aureus cellular proliferation protein #733.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS54422.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 12156; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   456 AA;

  Query Match             31.6%;  Score 1282;  DB 22;  Length 456;
  Best Local Similarity   54.4%;  Pred. No. 2.6e-105;
  Matches  257;  Conservative   64;  Mismatches  131;  Indels   20;  Gaps    5;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           |||:|||   ||:  ||    ||||| ||:|||||||||||||   ||:|||| | || |
Db       3 QTLFDKVWNRHVLYGKLGEPQLLYIDLHLIHEVTSPQAFEGLRLQNRKLRRPDLTFATLD 62

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
           |||||       || : ||:: :  |  ||::|  :|||  | :    |||||::||| |
Db      63 HNVPTI------DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDEQGIVHMVGPETG 115

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           | || |:|||||||:|||||||:|||||||||||| ||| |   : ||::| ::| |
Db     116 LTQPGKTIVCGDSHTATHGAFGAIAFGIGTSEVEHVFATQTLWQTKPKNLKIDINGTLPT 175

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           || :||::|| |   |   |||  :|| |  |::|||: ||:||||:|||||: |:: ||
Db     176 GVYAKDIILHLIKTYGVDFGTGYALEFTGETIKNLSMDGRMTICNMAIEGGAKYGIIQPD 235

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           :|||||:|||| |  :      |:   |: | ||  | :|  : :|   : | :||||:|
Db     236 DITFEYVKGRPFADNF-----AKSVDKWRELYSDDDAIFDRVIELDVSTLEPQVTWGTNP 290

Qy     309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
           |  |  :   |       |   : :|   ||||: |   ||| :  ||:|||||:|: |
Db     291 EMGVNFSEPFP-------EISDINDQRAYDYMGLEPGQKAEDIDLGYVFLGSCTNARLSD 343

Qy     369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
           |  |: :||| |  ||: :|:|||||  || :||: ||| ||: ||||||| ||||||||
Db     344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKREAEKLGLDTIFKNAGFEWREPGCSMCLGM 402

Qy     429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
           ||| :     |||||||||||||| | ||||:|| |||||  | ||  ||||:
Db     403 NPDQVPEGVHCASTSNRNFEGRQGKGARTHLVSPAMAAAAAIHGKFVDVRKV 454


RESULT 11
AAU37335
ID   AAU37335 standard; Protein; 456 AA.
XX
AC   AAU37335;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #1505.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS55194.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 12928; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
```

CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 456 AA;

Query Match 31.6%; Score 1282; DB 22; Length 456;
Best Local Similarity 54.4%; Pred. No. 2.6e-105;
Matches 257; Conservative 64; Mismatches 131; Indels 20; Gaps 5;

Qy 9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
Db 3 QTLFDKVWNRHVLYGKLGEPQLLYIDLHLIHEVTSPQAFEGLRLQNRKLRRPDLTFATLD 62

Qy 69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
Db 63 HNVPTI------DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDEQGIVHMVGPETG 115

Qy 129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
Db 116 LTQPGKTIVCGDSHTATHGAFGAIAFGIGTSEVEHVFATQTLWQTKPKNLKIDINGTLPT 175

Qy 189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
Db 176 GVYAKDIILHLIKTYGVDFGTGYALEFTGETIKNLSMDGRMTICNMAIEGGAKYGIIQPD 235

Qy 249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
Db 236 DITFEYVKGRPFADNF-----AKSVDKWRELYSDDDAIFDRVIELDVSTLEPQVTWGTNP 290

Qy 309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
Db 291 EMGVNFSEPFP-------EINDINDQRAYDYMGLEPGQKAEDIDLGYVFLGSCTNARLSD 343

Qy 369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
Db 344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKKEAEKLGLDTIFKNAGFEWREPGCSMCLGM 402

Qy 429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480
Db 403 NPDQVPEGVHCASTSNRNFEGRQGKGARTHLVSPAMAAAAAIHGKFVDVRKV 454

RESULT 12
ABB48172
ID ABB48172 standard; Protein; 462 AA.
XX
AC ABB48172;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #876.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 877; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 462 AA;

Query Match 31.6%; Score 1279.5; DB 23; Length 462;
Best Local Similarity 53.8%; Pred. No. 4.5e-105;
Matches 254; Conservative 65; Mismatches 134; Indels 19; Gaps 4;

Qy 9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
Db 3 KTLFDKLWNRHVIYGKEGEPQLLYVDLHLIHEVTSPQAFEGLRLENRPLRRPDKTFATMD 62

Qy 69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
Db 63 HNVPTEDIFNIQDLV-------AKKQIEALQTNCAEFGVTLADMGSDRQGIVHMVGPETG 115

Qy 129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
Db 116 LTQPGKVIVCGDSHTATHGAFGAIGFGIGSSEVEHVFATQTIWQQKPKSMGIEIINGKLPK 175

Qy 189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
Db 176 GVYAKDIILHLIATYGVAFGTGYAVEYYGETIRNMSMEERMTICNMAIEGGAKMGMMAPD 235

Qy 249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
Db 236 ETTFEYVRGREYAP----TDMDKAISDWKTLQTDSDAEYDLHIKMDASILEPYVTWGTNP 291

Qy 309 EDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIED 368
Db 292 EMGVPFSKAFP-------EIKDMNYERAYEYMGLKPGQTAEQIELGYVFIGSCTNARLSD 344

Qy 369 LRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLGM 428
Db 345 LEEAARIVKGNKVKNNIR-ALVVPGSRQVRNAAESIGLDKIFIEAGFEWREPGCSMCLGM 403

Qy 429 NPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL 480

```
Db    404 NPDQVPDGVHCASTSNRNFEGRQGKGARTHLVSPAMAAAAAINGHFIDIRKV 455


RESULT 13
AAR54216
ID   AAR54216 standard; Protein; 460 AA.
XX
AC   AAR54216;
XX
DT   09-NOV-1994  (first entry)
XX
DE   L.lactis branched amino acid synthesis leuC gene product.
XX
KW   branched amino acid; ilv operon; leucine; isoleucine; valine;
KW   biosynthesis; alpha-acetolactate synthase; diacetyl; food flavouring;
KW   attenuation; anti-terminator; Lactococcus.
XX
OS   Lactococcus lactis (subsp. lactis).
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 436
FT                   /note= "Val residue corresponds to CTG codon"
XX
PN   FR2696190-A.
XX
PD   01-APR-1994.
XX
PF   25-SEP-1992;   92FR-0011470.
XX
PR   25-SEP-1992;   92FR-0011470.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (AGRI-) AGRIC & FOOD RES COUNCIL.
XX
PI   Ehrlich S,  Godon J,  Renault P;
XX
DR   WPI; 1994-128287/16.
DR   N-PSDB; AAQ64211.
XX
PT   DNA coding for alpha-aceto:lactate synthase - for enhancing
PT   di:acetyl prodn. in microorganisms, esp. for mfr. of dairy prods.
XX
PS   Disclosure; Fig 2; 45pp; French.
XX
CC   The genes involved in the pathway for synthesis of branched amino
CC   acids in L.lactis subsp. lactis are organised in two units
CC   containing the leu (including leuC) and ilv genes, respectively.
CC   Both units are necessary for the synthesis of leucine but only the
CC   second unit is required for synthesis of Ile and Val. The ilvB
CC   and ilvN genes and the subunits of alpha-acetolactate synthase
CC   that they code for are claimed.
XX
SQ   Sequence   460 AA;

  Query Match             28.1%;  Score 1138.5;  DB 15;  Length 460;
  Best Local Similarity   50.4%;  Pred. No. 1.7e-92;
  Matches  240;  Conservative   61;  Mismatches  154;  Indels   21;  Gaps    6;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           :|::||:   ||:        ||||| |::||||||||:||| |||:||| | |  | |
Db       4 KTIFDKLWDQHVIAGNEGEPQLLYIDLHVIHEVTSPQAFQGLREAGRRVRRKDLTYGTLD 63

Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
          ||||| :   ::|:        |: |  |  :||||| |       | |||||:: || |
Db      64 HNVPTQNIFNIQDLI-------SKKQIDTFTKNVKEFDVPAETHGGKGQGIVHMVAPESG 116

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           | || |:|||||||:|:|||||:|||||||||||||||| :   : | |:|:  |
Db     117 RTQPGKTIVCGDSHTATNGAFGAIAFGIGTSEVEHVLATQTIWQVKPKRMKIEFQGHPQK 176

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           |: ||| :|  |   |   | |  :|: |  |  |||| ||:|||||| ||: |:: ||
Db     177 GIYSKDFILALIAKYGVDAGVGYAVEYSGDAISDLSMEERMTICNMSIEFGAKIGLMNPD 236

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           | |::|:|||  |||     : :|   |: | ||  |:||  : :|   : | :||||:|
Db     237 EKTYDYVKGREHAPK----NFDEAVSKWEKLVSDSDAQYDKILSLDVSQLKPMVTWGTNP 292

Qy     309 EDVVPITGV-VPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIE 367
                 |: :   | |       :  |  ||| || |    || :  :|||||||:|:
Db     293 -------GMGLEFGEKFPEINNDLNYERAYQYMDLKPGQTASDIDLGYIFIGSCTNARLG 345

Qy     368 DLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLG 427
           ||  || ::  |  |  : : :|||||  ||  || :||||||:||||||| ||| |||
Db     346 DLEEAAKIIGDRHIADGL-TGIVVPGSRPVKEAAEAQGLDKIFKEAGFEWREPGCSACLG 404

Qy     428 MNPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL-TD 482
           |||| :     |||||||||||||||   |||: || ||||| | ||  ||| | ||
Db     405 MNPDQIPEYVHCASTSNRNFEGRQGHNARTHVCSPAMAAAAAIAGKFVDVRMLVTD 460


RESULT 14
ABB54552
ID   ABB54552 standard; Protein; 460 AA.
XX
AC   ABB54552;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Lactococcus lactis protein leuC.
XX
KW   Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS   Lactococcus lactis IL1403.
XX
PN   FR2807446-A1.
XX
PD   12-OCT-2001.
XX
PF   11-APR-2000; 2000FR-0004630.
XX
PR   11-APR-2000; 2000FR-0004630.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Bolotine A,  Sorokine A,  Renault P,  Ehrlich SD;
XX
DR   WPI; 2002-043418/06.
XX
PT   New nucleotide sequence useful in the identification or Lactococcus
PT   lactis and related species -
XX
PS   Claim 6; SEQ ID No 1254; 2504pp; French.
XX
CC   The present invention is related to a Lactococcus lactis nucleotide
CC   sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC   nucleic acid sequence is useful in the detection and/or amplification of
CC   nucleic acid sequence, particularly to identify Lactococcus lactis or
CC   related species. The proteins of the invention are useful for the
CC   biosynthesis or biodegradation of a composition of interest. The
CC   invention helps research in lactic bacteria, particularly useful in the
CC   production of yogurt and cheese.
CC   Note: The sequence data for this patent is based on equivalent patent
CC   WO200177334 (published 18-OCT-2001) which is available in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   460 AA;

  Query Match             27.9%;  Score 1132.5;  DB 23;  Length 460;
  Best Local Similarity   50.4%;  Pred. No. 5.8e-92;
  Matches  240;  Conservative   59;  Mismatches  156;  Indels   21;  Gaps    6;

Qy       9 QTLYDKVLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTD 68
           :|::||:   ||:        ||||| |::||||||||:||| |||:||| | |  | |
Db       4 KTIFDKLWDQHVIAGNEGEPQLLYIDLHVIHEVTSPQAFQGLREAGRRVRRKDLTYGTLD 63
```

```
Qy      69 HNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQG 128
           |||||     ::|:        |: |  |  :||||| |       | |||||:: || |
Db      64 HNVPTQDIFNIQDLI-------SKKQIDTFTKNVKEFDVPAETHGGKGQGIVHMVAPESG 116

Qy     129 FTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAP 188
           | || |:|||||||:|:|||||:|||||||||||||||| :   : | |:|:  |
Db     117 RTQPGKTIVCGDSHTATNGAFGAIAFGIGTSEVEHVLATQTIWQVKPKRMKIEFQGHPQK 176

Qy     189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPD 248
           |: ||| :|  |   |   | |  :|: |  |  |||| ||:|||||| ||: |:: ||
Db     177 GIYSKDFILALIAKYGVDAGVGYAVEYSGDAISDLSMEERMTICNMSIEFGAKIGLMNPD 236

Qy     249 EITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGTSP 308
           | |::|:|||  |||     : :|   |: | ||  |:||  : :|   : | :||||:|
Db     237 EKTYDYVKGREHAPK----NFDEAVSKWEKLVSDSDAQYDKILSLDVSQLKPMVTWGTNP 292

Qy     309 EDVVPITGV-VPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRIE 367
                 |: :   | |       :  |  ||| || |     | :  :|||||||:|:
Db     293 -------GMGLEFGEKFPEINNDLNYERAYQYMDLKPGQTASAIDLGYIFIGSCTNARLG 345

Qy     368 DLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCLG 427
           ||  || ::  |  |  : : :|||||  ||  || :||||||:||||||| ||| |||
Db     346 DLEEAAKIIGDRHIADGL-TGIVVPGSRPVKEAAEAQGLDKIFKEAGFEWREPGCSACLG 404

Qy     428 MNPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVMAAAAGIVGKLADVRKL-TD 482
           ||||  :     |||||||||||||    |||| || ||||| | ||  ||| | ||
Db     405 MNPDQIPEYVHCASTSNRNFEGRQGHNARTHLCSPAMAAAAAIAGKFVDVRTLVTD 460
```

RESULT 15
AAW77717
ID AAW77717 standard; Protein; 264 AA.
XX
AC AAW77717;
XX
DT 30-OCT-1998 (first entry)
XX
DE 3-Isopropylmalate dehydratase protein.
XX
KW Staphylococcus aureus protein; immune response induction; eye infection;
KW antibody production; T-cell immune response; gastrointestinal infection;
KW respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW central nervous system; kidney infection; urinary tract infection;
KW antimicrobial compound identification; broad spectrum antibiotic;
KW therapy.
XX
OS Staphylococcus aureus.
XX
FH Key   Location/Qualifiers
FT Misc-difference 8
FT   /note= "unspecified, encoded by TNC"
FT Misc-difference 11
FT   /note= "unspecified, encoded by TNG"
XX
PN EP841394-A2.
XX
PD 13-MAY-1998.
XX
PF 24-SEP-1997; 97EP-0307485.
XX
PR 24-SEP-1996; 96US-0027032.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Black MT, Burnham MKR, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;
PI Ward JM;
XX
DR WPI; 1998-252940/23.
DR N-PSDB; AAV53507.
XX
PT New nucleic acid sequences from Staphylococcus aureus WCHU29 -
PT useful in vaccines and for treatment of bacterial infections of e.g.
PT respiratory tract and central nervous system
XX
PS Claim 11; Page 348-349; 390pp; English.
XX
CC This sequence represents a Staphylococcus aureus protein, that based on
CC homology with a Lactococcus lactis subsp lactis (Streptococcus
CC lactis) protein, is a 3-isopropylmalate dehydratase (Ec 4.2.1.33)
CC (isopropylmalate isomerase) (Alpha-Ipm isomerase) (Ipmi),
CC and is encoded by a DNA sequence of the invention.
CC The DNA sequences were isolated from Staphylococcus aureus WCHU29
CC (NCIMB 40771). Host cells containing the DNA sequences are used to
CC produce polypeptides or fragments. The proteins are used in the treatment
CC of disease, for inducing an immune response by administering them, to
CC produce antibody and/or T-cell immune response. Antagonists of the
CC proteins are used for the inhibition of bacterial polypeptides.
CC Conditions which may be treated include bacterial infections, especially
CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
CC urinary tract, skin, bones and joints. The proteins can also be used to
CC identify antimicrobial compounds which are broad spectrum antibiotics,
CC especially useful in the treatment of H. pylori infection.
XX
SQ Sequence 264 AA;

Query Match 18.5%; Score 749; DB 19; Length 264;
Best Local Similarity 53.8%; Pred. No. 3.9e-58;
Matches 147; Conservative 39; Mismatches 75; Indels 12; Gaps 3;

```
Qy      15 VLQAHVVDEKLDGTVLLYIDRHLVHEVTSPQAFEGLRNAGRKVRRPDCTLATTDHNVPTT 74
           |   ||:  |     ||||| ||:||||||||||||   ||:|||| | || ||||||
Db       1 VWNRHVLXGKXGDPQLLYIDLHLIHEVTSPQAFEGLRLQNRKLRRPDLTFATLDHNVPTI 60

Qy      75 SRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPEQGFTLPGT 134
                || : ||:: :  |  ||::|  :|||  | :    |||||::||| | | ||
Db      61 ------DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDEQGIVHMVGPETGLTQPGK 113

Qy     135 TVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGELAPGVSSKD 194
           |:|||||||:|||||||:|||||||||||| ||| |   : ||::| ::| |  || :||
Db     114 TIVCGDSHTATHGAFGAIAFGIGTSEVEHVFATQTLWQTKPKNLKIDINGTLPTGVYAKD 173

Qy     195 VVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVAPDEITFEY 254
           ::|| |   |   |||  :|| |  |::|||: ||:||||:|||||: |:: ||:|||||
Db     174 IILHLIKTYGVDFGTGYALEFTGETIKNLSMDGRMTICNMAIEGGAKYGIIQPDDITFEY 233

Qy     255 LKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKY 287
           :|||| |  :     |:   |: | ||   :|
Db     234 VKGRPFADNF-----AKSVDKWRELYSDGTTRY 261
```

RESULT 16
AAG81974
ID AAG81974 standard; Protein; 245 AA.
XX
AC AAG81974;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:1042.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.

```
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   N-PSDB; AAH52824.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 18; Page 303; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence   245 AA;

  Query Match             17.0%;  Score 688.5;  DB 22;  Length 245;
  Best Local Similarity   53.5%;  Pred. No. 8.6e-53;
  Matches  137;  Conservative   38;  Mismatches   68;  Indels   13;  Gaps    3;

Qy    225 MEARMSICNMSIEGGARAGMVAPDEITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPG 284
Db      1 MEARMTICNMAIEAGAKYGLMQPDETTFNYVKGRPYATDFDS-----SMAWWKELYSDDD 55

Qy    285 AKYDIDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKA 344
Db     56 AYFDKVIELDVTNLEPQVTWGTNPEMGVSFSNPFP-------EIKNANDQRAYDYMGLHP 108

Qy    345 GTPMEDIPVDKVFIGSCTNSRIEDLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEE 404
Db    109 GQKAEDIKLGYVFLGSCTNARLSDLIEASHIIKGQQVHPNI-TAIVVPGSRTVKKEAEAL 167

Qy    405 GLDKIFEEAGFEWREAGCSMCLGMNPDILAPQERCASTSNRNFEGRQGAGGRTHLMSPVM 464
Db    168 GLDKLFKDAGFEWREPGCSMCLGMNPDQVPEGVHCASTSNRNFEGRQGKGARTHLVSPAM 227

Qy    465 AAAAGIVGKLADVRKL 480
Db    228 AAAAAINGKFIDVRKV 243


RESULT 17
AAB96358
ID   AAB96358 standard; Protein; 423 AA.
XX
AC   AAB96358;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Putative 3-isopropylmalate dehydratase/aconitase large subunit #2.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS   Pyrococcus abyssi.
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 7; Pages 1038-1040; 1657pp; French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC   a hyperthermophilic archaeon, which is isolated from deep-sea
CC   hydrothermal vents. The present sequence is one such P. abyssi protein.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence   423 AA;

  Query Match             15.7%;  Score 637.5;  DB 22;  Length 423;
  Best Local Similarity   35.3%;  Pred. No. 7.3e-48;
  Matches  166;  Conservative   77;  Mismatches  166;  Indels   61;  Gaps   10;

Qy     10 TLYDKVLQAHVVDE--KLDGTVLLYIDRHLVHEVTSPQAFEGLRNAG-RKVRRPDCTLAT 66
Db      4 TIAEKILADHSEREEVKPGEIVMAKLDFVFGNDVTMPLAIKKFRELGVKRVFDRERIAIV 63

Qy     67 TDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRQGIVHVIGPE 126
Db     64 LDHFTPN------KDI-------KSAEQCKSSREFAKEMGIKWF-FEGGSVGVEHCLLPE 109

Qy    127 QGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNMRIQVDGEL 186
Db    110 LGLVLPGDLIIGADSHTCTYGALGAFATGVGSTDLAVAMATGEAWFRVPETMKFIYEGEL 169

Qy    187 APGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEGGARAGMVA 246
Db    170 QPYVTGKDLILHTIGDIGVNGALYKVMEFSGSVIEELSVEQRMTMSNMAIEAGAKTGIIE 229

Qy    247 PDEITFEYLKGRPLAPKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDAKDIVPTLTWGT 306
Db    230 PDKKTLDYVKER-------------AKRKFKVYKSDEDAKYYKVIEYDVTNWEPVVAFPH 276

Qy    307 SPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDKVFIGSCTNSRI 366
Db    277 LPENTVPI-----------SKAAKM-----------------NIKIDQVFIGSCTNGRI 307

Qy    367 EDLRAAAAVVKGRKKAPNVKSAMVVPGSGLVKTQAEEEGLDKIFEEAGFEWREAGCSMCL 426
Db    308 EDLRMAAEILEGQKVAKWVR-LIVIPCSPTVYWKALKEGLIEIFLEAGAVIGPPTCGPCL 366

Qy    427 GMNPDILAPQERCASTSNRNFEGRQG-AGGRTHLMSPVMAAAAGIVGKLA 475
Db    367 GGHMGVLASGERAVSTTNRNFVGRMGHPKSEVYLANPYVAAASAVLGRIA 416


RESULT 18
AAG29924
```

ID AAG29924 standard; Protein; 461 AA.
XX
AC AAG29924;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35684.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 558.5; DB 21; Length 461;
Best Local Similarity 32.1%; Pred. No. 9.4e-41;
Matches 164; Conservative 73; Mismatches 175; Indels 99; Gaps 22;

```
Qy      3 GAESTPQTLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKV 57
          |:  |  |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | |
Db     13 GSVKTGMTMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKA 70

Qy     58 R--RPDCTLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFG 111
          :   |:  :   || : |  ::| :::      |  | |   |:|:| |    : |
Db     71 KVWDPEKIVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFK 122

Qy    112 LSDKRQGIVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
          :   :|: ||   ::|   ||  ::  |||| | ||||  | ||| ::   || |  ::
Db    123 ANPDYKGVCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKIL 182

Qy    172 TKRSKNMRIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSI 231
            |   ||  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::
```

```
Db    183 LKVPPTMRFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTL 242

Qy    232 CNMSIEGGARAGMVAPDEITFEYLKGRPLA---PKYDSPEWHKATQYWKNLQSDPGAKYD 288
          ||| :| | : |:: ||  |  |::   |:   | |                ||  | :
Db    243 CNMVVEAGGKNGVIPPDATTLNYVEACILSCFLPVY----------------SDGNASFV 286

Qy    289 IDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPM 348
          |   |   : | :     |          ||    | | |
Db    287 ADYRFDVSKLEPVV---AKPHS--------PDNRALARECK------------------- 316

Qy    349 EDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------KSAMVVPGS 394
          |: :|:|:|||||  : ||  ||| :    ||| | |              |: |||:
Db    317 -DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDVYALPVPGA 375

Qy    395 GLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCASTSNRNFEG 449
          |  || |:      |||||| :      |  |||   |  |    ||  | ||:|||| |
Db    376 G-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVSTTNRNFPG 427

Qy    450 RQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
          | |   |: :| ||  |||: : |::|| |:
Db    428 RMGHKEGQIYLASPYTAAASALTGRVADPRE 458
```

RESULT 19
AAG40226
ID AAG40226 standard; Protein; 469 AA.
XX
AC AAG40226;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49883.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

```
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             13.8%;  Score 558.5;  DB 21;  Length 469;
  Best Local Similarity   31.6%;  Pred. No. 9.7e-41;
  Matches  164;  Conservative   72;  Mismatches  176;  Indels  107;  Gaps   22;

Qy       3 GAESTPQTLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKV 57
Db      13 GSVKTGMTMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKA 70

Qy      58 R--RPDCTLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFG 111
Db      71 KVWDPEKIVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFK 122

Qy     112 LSDKRQGIVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
Db     123 ANPDYKGVCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKIL 182

Qy     172 TKRSKNMRIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSI 231
Db     183 LKVPPTMRFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTL 242

Qy     232 CNMSIEGGARAGMVAPDEITFEYLK-----------GRPLAPKYDSPEWHKATQYWKNLQ 280
Db     243 CNMVVEAGGKNGVIPPDATTLNYVEACILSCFLNRTSVPFEPVY---------------- 286

Qy     281 SDPGAKYDIDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYM 340
Db     287 SDGNASFVADYRFDVSKLEPVV---AKPHS--------PDNRALARECK----------- 324

Qy     341 GLKAGTPMEDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------K 386
Db     325 ---------DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDV 375

Qy     387 SAMVVPGSGLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCAS 441
Db     376 YALPVPGAG-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVS 427

Qy     442 TSNRNFEGRQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
Db     428 TTNRNFPGRMGHKEGQIYLASPYTAAASALTGRVADPRE 466


RESULT 20
AAG29923
ID   AAG29923 standard; Protein; 509 AA.
XX
AC   AAG29923;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 35683.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 558.5; DB 21; Length 509;
Best Local Similarity 32.1%; Pred. No. 1.1e-40;
Matches 164; Conservative 73; Mismatches 175; Indels 99; Gaps 22;

```
Qy      3 GAESTPQTLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKV 57
          |:  |  |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | |
Db     61 GSVKTGMTMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKA 118

Qy     58 R--RPDCTLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFG 111
          :  |: :  || : |  ::| :::      |  |  |   |:|:| |     :  |
Db    119 KVWDPEKIVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFK 170

Qy    112 LSDKRQGIVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
          :   :|: ||   ::|   ||  ::  |||| | ||||  | ||| ::   || |  ::
Db    171 ANPDYKGVCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKIL 230

Qy    172 TKRSKNMRIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSI 231
          |   ||  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::
Db    231 LKVPPTMRFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTL 290

Qy    232 CNMSIEGGARAGMVAPDEITFEYLKGRPLA---PKYDSPEWHKATQYWKNLQSDPGAKYD 288
          ||| :| | : |:: ||  |  |::   |:   | |                ||  | :
Db    291 CNMVVEAGGKNGVIPPDATTLNYVEACILSCFLPVY---------------SDGNASFV 334

Qy    289 IDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPM 348
          |   |   : | :      |          ||    | | |
Db    335 ADYRFDVSKLEPVV---AKPHS--------PDNRALARECK------------------- 364

Qy    349 EDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------KSAMVVPGS 394
          |: :|:|:|||||  : ||  ||| :    ||| | |                |: |||:
Db    365 -DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDVYALPVPGA 423

Qy    395 GLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCASTSNRNFEG 449
          |  || |:      ||||| :      |  |||   | |    ||  | ||:|||| |
Db    424 G-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVSTTNRNFPG 475

Qy    450 RQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
          | |   |: :| ||  |||: : |::|| |:
Db    476 RMGHKEGQIYLASPYTAAASALTGRVADPRE 506
```

```
RESULT 21
ABB92924
ID   ABB92924 standard; Protein; 509 AA.
XX
AC   ABB92924;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 2135.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 2135; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   509 AA;

  Query Match             13.8%;  Score 558.5;  DB 23;  Length 509;
  Best Local Similarity   32.1%;  Pred. No. 1.1e-40;
  Matches  164;  Conservative   73;  Mismatches  175;  Indels   99;  Gaps   22;

Qy       3 GAESTPQTLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKV 57
           |:  |  |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | |
Db      61 GSVKTGMTMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKA 118

Qy      58 R--RPDCTLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFG 111
           :   |:  :   || : |  ::| :::      |  |  |   |:|:| |     :  |
Db     119 KVWDPEKIVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFK 170

Qy     112 LSDKRQGIVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
            :   :|: ||   ::|   ||  ::  |||| | ||||  | ||| ::   || |  ::
Db     171 ANPDYKGVCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKIL 230

Qy     172 TKRSKNMRIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSI 231
           |    ||  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::
Db     231 LKVPPTMRFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTL 290

Qy     232 CNMSIEGGARAGMVAPDEITFEYLKGRPLA---PKYDSPEWHKATQYWKNLQSDPGAKYD 288
           ||| :| | : |:: ||  |  |::   |:   | |                ||  | :
Db     291 CNMVVEAGGKNGVIPPDATTLNYVEACILSCFLPVY---------------SDGNASFV 334

Qy     289 IDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPM 348
           |   |   : | :     |          ||    | | |

Db     335 ADYRFDVSKLEPVV---AKPHS--------PDNRALARECK------------------- 364

Qy     349 EDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------KSAMVVPGS 394
           |: :|:|:|||||  : ||  ||| :    ||| | |              |: |||:
Db     365 -DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDVYALPVPGA 423

Qy     395 GLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCASTSNRNFEG 449
           |  || |:       |||||| :       |  |||   |  |    ||  | ||:|||| |
Db     424 G-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVSTTNRNFPG 475

Qy     450 RQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
           | |   |: :| ||  |||: : |::|| |:
Db     476 RMGHKEGQIYLASPYTAAASALTGRVADPRE 506


RESULT 22
AAG40225
ID   AAG40225 standard; Protein; 517 AA.
XX
AC   AAG40225;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 49882.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
```

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 558.5; DB 21; Length 517;

```
  Best Local Similarity   31.6%;  Pred. No. 1.1e-40;
  Matches  164;  Conservative   72;  Mismatches  176;  Indels  107;  Gaps   22;

Qy       3 GAESTPQTLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKV 57
           |:  |  |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | |
Db      61 GSVKTGMTMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKA 118

Qy      58 R--RPDCTLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFG 111
           :   |:  :   || : |  ::| :::      |  |  |   |:|:| |     :  |
Db     119 KVWDPEKIVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFK 170

Qy     112 LSDKRQGIVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLI 171
           :   :|: ||   ::|   ||  ::  |||| | ||||  | ||| ::   || |  ::
Db     171 ANPDYKGVCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKIL 230

Qy     172 TKRSKNMRIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSI 231
           |    ||  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::
Db     231 LKVPPTMRFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTL 290

Qy     232 CNMSIEGGARAGMVAPDEITFEYLK-----------GRPLAPKYDSPEWHKATQYWKNLQ 280
           ||| :| | : |:: ||  |  |::            |  | |
Db     291 CNMVVEAGGKNGVIPPDATTLNYVEACILSCFLNRTSVPFEPVY---------------- 334

Qy     281 SDPGAKYDIDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYM 340
           ||  | :  |   |   : | :     |          ||    | | |
Db     335 SDGNASFVADYRFDVSKLEPVV---AKPHS--------PDNRALARECK----------- 372

Qy     341 GLKAGTPMEDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------K 386
                 |: :|:|:|||||  : ||  ||| :    ||| | |
Db     373 ---------DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDV 423

Qy     387 SAMVVPGSGLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCAS 441
           |: |||:|  || |:      ||||| :      |  |||   |  |   ||  | |
Db     424 YALPVPGAG-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVS 475

Qy     442 TSNRNFEGRQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
           |:|||| || |   |: :| ||  |||: : |::|| |:
Db     476 TTNRNFPGRMGHKEGQIYLASPYTAAASALTGRVADPRE 514


RESULT 23
AAG29925
ID   AAG29925 standard; Protein; 443 AA.
XX
AC   AAG29925;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 35685.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
```

```
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             13.6%;  Score 550.5;  DB 21;  Length 443;
  Best Local Similarity   32.1%;  Pred. No. 4.6e-40;
  Matches  162;  Conservative   72;  Mismatches  171;  Indels   99;  Gaps   22;

Qy      10 TLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKVR--RPDC 62
           |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | | :   |:
Db       2 TMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKAKVWDPEK 59

Qy      63 TLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFGLSDKRQG 118
             :  || : |  ::| :::      |  |  |   |:|:| |     :  |  :   :|
Db      60 IVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFKANPDYKG 111

Qy     119 IVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNM 178
           : ||   ::|   ||  ::  |||| | ||||  | ||| ::   || |  :: |    |
Db     112 VCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKILLKVPPTM 171

Qy     179 RIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEG 238
           |  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::||| :|
Db     172 RFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTLCNMVVEA 231

Qy     239 GARAGMVAPDEITFEYLKGRPLA---PKYDSPEWHKATQYWKNLQSDPGAKYDIDVFIDA 295
           | : |:: ||  |  |::   |:   | |              ||  | :  |   |
Db     232 GGKNGVIPPDATTLNYVEACILSCFLPVY---------------SDGNASFVADYRFDV 275

Qy     296 KDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTPMEDIPVDK 355
             : | :    |          ||    | | |                    |: :|:
Db     276 SKLEPVV---AKPHS--------PDNRALARECK-------------------DVKIDR 304

Qy     356 VFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------KSAMVVPGSGLVKTQA 401
           |:|||||  : ||  ||| :    ||| | |               |: |||:|  || |
Db     305 VYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDVYALPVPGAG-GKTCA 363

Qy     402 EEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCASTSNRNFEGRQG-AGG 455
           :      |||||| :      |  |||   |  |    ||  | ||:|||| || |   |
Db     364 Q------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVSTTNRNFPGRMGHKEG 416

Qy     456 RTHLMSPVMAAAAGIVGKLADVRK 479
           : :| ||  |||: : |::|| |:
Db     417 QIYLASPYTAAASALTGRVADPRE 440


RESULT 24
AAG40227
ID   AAG40227 standard; Protein; 451 AA.
XX
AC   AAG40227;
XX
DT   18-OCT-2000  (first entry)
XX
```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 49884.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 550.5; DB 21; Length 451;
Best Local Similarity 31.6%; Pred. No. 4.7e-40;
Matches 162; Conservative 71; Mismatches 172; Indels 107; Gaps 22;

```
Qy      10 TLYDKVLQAHVVDEKL---DGTVLLYIDRHLVHEVTSPQAFEGL--RNAGRKVR--RPDC 62
           |: :|:| |   :: |      : : :|  : |:|  | || |:  |  | | :   |:
Db       2 TMTEKIL-ARASEKSLVVPGDNIWVNVDVLMTHDVCGPGAF-GIFKREFGEKAKVWDPEK 59

Qy      63 TLATTDHNVPTTSRKALKDIASFIKEDDSRTQCVTLEENVKEF----GVTYFGLSDKRQG 118
              :  || : |  ::| :::      | | |   |:|:| |     :  |  :   :|
Db      60 IVVIPDHYIFTADKRANRNV------DIMREHC--REQNIKYFYDITDLGNFKANPDYKG 111

Qy     119 IVHVIGPEQGFTLPGTTVVCGDSHTSTHGAFGALAFGIGTSEVEHVLATQCLITKRSKNM 178
           : ||   ::|   ||  ::  |||| | |||| | ||| ::   || |  :: |     |
Db     112 VCHVALAQEGHCRPGEVLLGTDSHTCTAGAFGQFATGIGNTDAGFVLGTGKILLKVPPTM 171

Qy     179 RIQVDGELAPGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMEARMSICNMSIEG 238
           |  :|||:   : :||::|  || |  || |   :|| |: | ||||| ||::||| :|
Db     172 RFILDGEMPSYLQAKDLILQIIGEISVAGATYKTMEFSGTTIESLSMEERMTLCNMVVEA 231

Qy     239 GARAGMVAPDEITFEYLK----------GRPLAPKYDSPEWHKATQYWKNLQSDPGAKY 287
           | : |:: ||  |  |::             | | |                ||  | :
Db     232 GGKNGVIPPDATTLNYVEACILSCFLNRTSVPFEPVY----------SDGNASF 275

Qy     288 DIDVFIDAKDIVPTLTWGTSPEDVVPITGVVPDPETFATEAKKADGRRMLQYMGLKAGTP 347
           |   |   : | :    |          ||    | | |
Db     276 VADYRFDVSKLEPVV---AKPHS--------PDNRALARECK---------- 306

Qy     348 MEDIPVDKVFIGSCTNSRIEDLRAAAAV--VKGRK-KAPNV----------KSAMVVPG 393
           |: :|:|:|||||  : ||  ||| :    ||| | |                |: |||
Db     307 --DVKIDRVYIGSCTGGKTEDFMAAAKLFHAAGRKVKVPTFLVPATQKVWMDVYALPVPG 364

Qy     394 SGLVKTQAEEEGLDKIFEEAGFEW-REAGCSMCLGMNPDILA----PQERCASTSNRNFE 448
           :|  || |:       |||||| :       |  |||   |  |    ||  | ||:||||
Db     365 AG-GKTCAQ------IFEEAGCDTPASPSCGACLGGPADTYARLNEPQV-CVSTTNRNFP 416

Qy     449 GRQG-AGGRTHLMSPVMAAAAGIVGKLADVRK 479
           || |   |: :| ||  |||: : |::|| |:
Db     417 GRMGHKEGQIYLASPYTAAASALTGRVADPRE 448
```

RESULT 25
AAU36307
ID AAU36307 standard; Protein; 212 AA.
XX
AC AAU36307;
XX
DT 14-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa cellular proliferation protein #297.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS54166.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 11900; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an

```
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   212 AA;

  Query Match            13.4%;  Score 542;  DB 22;  Length 212;
  Best Local Similarity  53.3%;  Pred. No. 8e-40;
  Matches  112;  Conservative   34;  Mismatches   50;  Indels   14;  Gaps    6;

Qy     543 FTILKGIAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRF---NEDGTEKS---- 595
           :|   |: |||::|||||| |||||||:|||| |  || | |:    : | : |
Db       4 YTQTTGLVAPLDRANVDTDQIIPKQFLKSIKRTGFGPNLFDEWRYLDVGQPGQDNSKRPL 63

Qy     596 --DFVLNKEPYRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIAPSFADIFFNNSFK 653
             |||||:  |: ||:|:    |||||||||||||||:::| |:|||||:||||||||||
Db      64 NPDFVLNQPRYQGASVLLAR-ENFGCGSSREHAPWALDEYGFRTVIAPSYADIFFNNSFK 122

Qy     654 NGMLPIPIKDQAQIEAI--AAEARAGKEIEVDLPNQLIKNATGETICTFEVEEFRKHCLV 711
           ||:||| |  :|::: :    ||  | :: :||  | :    |: :  |||: ||||||:
Db     123 NGLLPI-ILPEAEVDELFRQVEANEGYQLSIDLAAQTVTRPDGK-VLGFEVDPFRKHCLL 180

Qy     712 NGLDDIGLTMQMEDKIAEFEAKMTRETPWL 741
           |||||||||:|  | |  ||    :: |||
Db     181 NGLDDIGLTLQDADAIRAFEDGYRQQQPWL 210


Search completed: March 17, 2003, 08:49:07
Job time : 53 secs
```